# 5 Zufälliges, Wahrscheinliches und Normales

In diesem Kapitel wollen wir uns den *Zufall* ein wenig näher anschauen und auch die Theorie der damit im Zusammenhang stehenden *Wahrscheinlichkeit* für das Eintreffen eines Ereignisses, das wir als «zufällig» einstufen.

## 5.1 Zufall

Beim Finale der UEFA Champions League am 19. Mai 2012 in München spielten in der Mannschaft des Chelsea FC gleich zwei «Geburtstagspaare», also je zwei Spieler, die am selben Tag Geburtstag feiern: Salomon Kalou und Ryan Bertrand (5. August) sowie David Luiz und John Obi Mikel (22. April). Die beiden letztgenannten sind sogar nicht nur am selben Tag sondern auch im selben Jahr (1987) auf die Welt gekommen.
Inklusive des Schiedsrichters und der (tatsächlich zum Einsatz gekommenen) Ersatzspieler waren 27 Personen am Spielfeld tätig. Würdest du – hätte ich es nicht bereits verraten – darauf wetten, dass mindestens zwei von ihnen am selben Tag Geburtstag haben?

So ein *Zufall*: Geburtstags(zu)fälle: Intuitiv empfinden wir das als außergewöhnliches Zusammentreffen – und als *großen Zufall*.

Der Begriff **Zufall** beim Beobachten einer *Zufalls*größe soll unterstreichen, dass das Ergebnis dieser Beobachtung nicht vorhersehbar oder *deterministisch*[1] ist.

[1] vom lat. *determinare* = bestimmen, festsetzen. Ein «deterministisches» Ergebnis bedeutet: Es gibt einen funktionalen Zusammenhang zwischen den Eingangsparametern und dem Ergebnis und wenn wir alle Eingangsparameter kennen, kennen wir auch das Ergebnis.

Wenn wir zum Beispiel eine Münze werfen, wissen wir nicht, ob wir Kopf oder Zahl erhalten. Wenn wir auf der Hauptseite von *de.wikipedia.org* auf den Link «Zufälliger Artikel» klicken (oder gleichzeitig die Tasten `alt-shift-x`), rufen wir irgendeinen, zufällig ausgewählten Artikel auf. Alles zufällige Ereignisse. Aber auch andere, kompliziertere Dinge wie der Wert eines Aktienindex oder die Paketumlaufzeit in einem IP-Netzwerk sind – im Sinne der Statistik – zufällige Ereignisse.

Trotz aller Zufälligkeit würden wir gerne herausfinden, ob es nicht doch irgendwelche «Gesetzmäßigkeiten» gibt, wie wir dem Zufall auf die Schliche kommen und ihn in den Griff bekommen könnten.

Wenn wir die Geburtstage von Fußballspielern «beobachten», ist das aus statistischer Sicht dasselbe, wie wenn wir mit einem 365-seitigem Würfel würfeln oder aus einer Urne mit 365 verschiedenen Kugeln (mit jeweils einem aufgedruckten Tagesdatum) eine beliebige Kugel herausziehen. Beides – «Würfeln» und «Ziehen aus einer Urne» – wird oft als anschauliches Denkmodell für ein *Zufallsexperiment* (auch: *zufälliger Versuch*) verwendet. Ein **Zufallsexperiment** ist ein Vorgang, der – zumindest im Prinzip – beliebig oft wiederholbar ist und die jeweiligen Ergebnisse sind innerhalb einer Menge möglicher Ausgänge ungewiss, eben zufällig. Das Ergebnis eines zufälligen Versuches bezeichnen wir dann als ein **Zufallsereignis** $E$.

Der einzelne Wert, den die Zufallsgröße nach der Beobachtung (als Ergebnis des Zufallsexperiments) annimmt, ist die **Realisierung** $x$ der Zufallsgröße $X$. Beim Zufallsexperiment «Würfeln» und Beobachtung der Zufallsvariable $X = Augenzahl$ können wir zum Beispiel im 3. Versuch die Zahl 4 erhalten. $x_3 = 4$ ist dann eine Realisierung der Zufallsvariable $Augenzahl$. Ein anderes Beispiel für die Realisierung einer Zufallsgröße ist der tägliche Schlusskurs des *Dow Jones Industrial Average*, nämlich der Zufallsgröße «Kursindex der dreißig größten US-Unternehmen am Ende eines Börsentages an der New York Stock Exchange».

Realisierungen von Zufallsgrößen sind selbst übrigens nicht mehr zufällig. Sie haben ja einen bestimmten Wert.

In der Alltagssprache bezeichnen wir meist Ereignisse, die ohne offensichtlichen Grund eintreten, und deren Eintritt uns als ziemlich unwahrscheinlich scheint, als «zufällig». Zum Beispiel wenn in einer Fußballmannschaft zwei Spieler am selben Tag Geburtstag haben, oder wenn man sechsmal hintereinander würfelt und dabei genau die Abfolge $1, 2, 3, 4, 5, 6$ erhält. Tatsächlich ist aber zum Beispiel auch die Abfolge $3, 3, 5, 2, 6, 1$ genauso «zufällig».

Um den Zufall mathematisch-statistisch beschreiben und modellieren zu können, benötigen wir die Wahrscheinlichkeitsrechnung.

Zur erstmaligen Verwendung des Wortes «Lockdown» kam es laut dem deutschen Sprachwissenschafter Anatol Stefanowitsch bei der Beschreibung eines Vorfalls in einem US-Gefängnis 1973: Am 6. Dezember dieses Jahres wurde durch mehrere Mithäftlinge 32 mal auf einen Häftling eingestochen, der einen «Kollegen» unabsichtlich angefahren und sich dafür nicht entschuldigt hatte. Daraufhin wurden zunächst alle Gefangenen für einige Zeit in ihren Zellen eingesperrt, was als «Lockdown» bezeichnet wurde.
Wirklich spooky an der Sache aber: Der Gefangene, der Opfer dieser Messerattacke wurde, hieß Juan Vallejo *Corona*.

Was 1973 aus Sicht der restlichen Welt vermutlich nur eine unbedeutende Episode war, klingt beinahe 50 Jahre später nach einem aberwitzigen Zufall.

## 5.2 Ein bisschen Wahrscheinlichkeitsrechnung

Was hinter einer «Wahrscheinlichkeit» steckt, davon hat vermutlich jeder seine subjektive Vorstellung. Man überlegt beispielsweise, wie wahrscheinlich es ist, dass man durch die Abschlussprüfungen am Ende des Semesters kommt, dass im Urlaub im Salzkammergut eine Woche lang die Sonne scheint, oder dass man im Lotto gewinnt. Und obwohl die Wahrscheinlichkeit für Letzteres wirklich sehr klein ist[2], gibt es mehr Lottospieler als Urlauber im Salzkammergut.

Auch beim Festlegen von Börsenstrategien und Devisengeschäften rechnet man – bewusst oder unbewusst – mit Wahrscheinlichkeiten (dort nennt man es «Spekulieren»), aber auch wenn man die Erfolgschancen für ein neues Produkt einschätzt oder bei anderen Marketingentscheidungen. Die Wahrscheinlichkeitsrechnung spielt in der Versicherungsmathematik eine Rolle, bei der Qualitätskontrolle, bei der Optimierung von Produktion und Lagerhaltung und so fort. Trotz ihrer Wichtigkeit ist sie bei Wirtschafts- und Informatikstudierenden eher unbeliebt (insbesondere als Prüfungsgegenstand), was vielleicht auch daran liegt, dass es mehrere Definitionen für sie gibt. Die «subjektive Wahrscheinlichkeitsdefinition» haben wir bereits angesprochen, es gibt aber natürlich auch eine «richtige» Definition:

Sie stammt von *Laplace*[3] und gibt folgendes Verhältnis wieder:

$$P(E) = \frac{\textit{Zahl der möglichen Eintrittsfälle von E}}{\textit{Gesamtzahl aller überhaupt möglichen Ausgänge}} \tag{5.1}$$

[2] Die Chance auf einen 6er bei «6 aus 45» beträgt 1:8.145.060. Dennoch hat 2008 eine Kärntnerin bei einem Sechser rund 1.7 Millionen Euro gewonnen, nachdem ihr Ehemann bereits 1999 einen Sechser mit umgerechnet rund 1 Million Euro getippt hatte.

[3] *Pierre-Simon Marquis de Laplace*, frz. Mathematiker, Astronom und Physiker, 1749-1827

In *MS Excel* und *LibreOffice Calc* können wir auf zwei Arten Zufallszahlen erzeugen:
Die Funktion `=ZUFALLSZAHL()` gibt eine reelle Zufallszahl größer oder gleich 0 und kleiner als 1 zurück.
Mit `=ZUFALLSBEREICH(UntereZahl; ObereZahl)` erhalten wir eine *ganze* Zahl, die (zufällig) irgendwo zwischen den beiden angegebenen Grenzen liegt. (Dabei sind die untere und obere Grenze in den möglichen Zahlen ebenfalls inkludiert). Die Zufallszahl kann (bei entsprechender Angabe der beiden Grenzen) auch negativ sein, ist aber in jedem Fall eine ganze Zahl.
Will man eine reelle Zufallszahl zwischen den Grenzen $a$ und $b$ haben, muss man ein wenig kreativ sein und angeben: `=ZUFALLSZAHL()*(b-a)+a`.

In *R* gibt es mehrere (und «ausgefeiltere») Methoden, um Zufallszahlen zu erzeugen. Mit `sample(10:20,2)` erhalten wir zum Beispiel 2 zufällige Zahlen aus dem Intervall von 10 bis 20. Weitere Funktionen sind zum Beispiel `rnorm` oder `runif`, zu deren Bedeutung wir später kommen.

Falls du zufällig einmal eine Zufallszahl brauchst: Der PC kann aushelfen. Oder das Internet unter *www.random.org* oder *www.randomnumbers.info*

$P(E)$ ist die **Wahrscheinlichkeit**[4] für das Eintreten des Zufallsereignisses $E$.

Formel 5.1 sieht sehr einfach aus und wir können gleich ein Beispiel rechnen:

**Beispiel 22** *Bei der Fußball-Europameisterschaft 1968 in Italien wurde nach einem 0:0 im Halbfinale zwischen Italien und der UdSSR per Münzwurf entschieden, dass Italien ins Endspiel aufstieg (Das es schließlich auch gewann). Und auch bei der Fußball-EM der Frauen 2013 in Schweden musste nach den Vorrundenspielen per Münzwurf entschieden werden, ob Dänemark oder Russland ins Viertelfinale einziehen wird (Der Zufall brachte Dänemark Glück).*

*Unter der Annahme, dass sich die Münze bei einem Losentscheid nicht buchstäblich in Luft auflöst oder in ein Erdloch fällt (und auch nicht auf der Kante zu stehen kommt[5]): Wie groß ist bei einem Münzwurf die Wahrscheinlichkeit für «Kopf»?*

*Es gibt 2 mögliche Ausgänge (nämlich «Kopf» oder «Zahl»), und einen Eintrittsfall von*

[4] Das $P$ kommt vom lat. *probabilitas* = Wahrscheinlichkeit
[5] wie zum Beispiel 1965 beim Europapokal-Entscheidungsspiel zwischen dem 1. FC Köln und dem FC Liverpool, siehe https://youtu.be/-P8zft3Yc1M

Im 17. Jhdt. wurde *Blaise Pascal* (frz. Mathematiker, 1623-1662) vom frz. Schriftsteller und Berufsspieler *Antoine Gombaud Chevalier de Méré* (1607-1684) mit der Frage konfrontiert, wie der Einsatz bei einem bestimmten Würfelspiel fairerweise aufzuteilen ist, wenn das Spiel vorzeitig abgebrochen werden muss. Es ging also um die Frage nach der Wahrscheinlichkeit, mit der jeder Teilnehmer das Spiel gewinnen würde, wenn es fortgesetzt werden würde. Pascal beriet sich daraufhin in mehreren Briefwechseln mit seinem Kollegen *Pierre de Fermat* (frz. Mathematiker und Jurist, 1607-1665). Damit war die *Wahrscheinlichkeitsrechnung* geboren.

Zurück zum Ursprung: Die Anfänge der Wahrscheinlichkeitsrechnung

In Vernon River-Stratford, einem Wahlbezirk in der ostkanadischen Provinz Prince Edward Island, erreichten am 4.5.2015 Mary Ellen McInnis und Alan McIsaac jeweils 1.173 Stimmen. Daraufhin wurde per Münzwurf entschieden, wer den Sitz im Provinzparlament erhalten sollte.
Dabei sieht die Wahlordnung vor, dass sich aus der alphabetischen Reihenfolge der Nachnamen ergibt, welchem Kandidaten «Kopf» und welchem «Zahl» zugeordnet wird. Und selbst da wurde der Zufall ziemlich herausgefordert. Unterscheiden sich die beiden Namen McInnis und McIsaac doch erst ab dem vierten Buchstaben.
Letztlich fiel die Münze auf «Zahl» und bescherte somit Alan McIsaac das Mandat.

Der Zufall gibt sich manchmal auch demokratisch.

*E (nämlich «Kopf»). Somit:*

$$P(E) = \frac{\textit{Eintrittsfälle}}{\textit{Ausgangsmöglichkeiten}} = \frac{1}{2} = 0.5 = 50\%$$

**Aufgabe 16** In Oberösterreich gibt es insgesamt 6 630 Orte. 40 davon heißen «Au». Angenommen alle 6 630 Orte werden auf Kärtchen geschrieben und daraus eine beliebige Karte herausgezogen. Wie groß ist die Wahrscheinlichkeit, dass darauf *Au* steht?

**Aufgabe 17** Unter der Annahme, dass beim Würfeln jede Augenzahl gleich wahrscheinlich ist: Wie groß ist die Wahrscheinlichkeit, eine gerade Zahl zu würfeln?

In der Bulgarischen Lotterie, in der man 6 aus 49 Zahlen tippt, wurden am 6. September 2009 und bei der darauffolgenden Ziehung am 10. September 2009 exakt dieselben Zahlen gezogen (4, 15, 23, 24, 35, und 42). Mag das schon unglaublich genug sein, ist es mindestens ebenso unerwartet, dass beim zweiten mal gleich 18 Spieler einen Sechser erzielten, also entgegen jeglicher landläufiger Vorstellung von Wahrscheinlichkeiten dieselben Zahlen wie bei der vorherigen Ziehung getippt hatten. (Für ihren Sechser gewannen sie umgerechnet jeweils ca. 5 200 EUR).

Auch Ereignisse, die nur mit einer Wahrscheinlichkeit von 1:4.2 Millionen auftreten, treten unwahrscheinlicherweise irgendwann auf.

Neben einer Definition benötigen wir für das Berechnen von und Rechnen mit Wahrscheinlichkeiten noch Rechenregeln. Die einfachsten[6] sind:

$$P(E) + P(not\ E) = 1 \tag{5.2}$$
$$P(not\ E) = 1 - P(E) \tag{5.3}$$
$$P(E_1\ and\ E_2) = P(E_1) \cdot P(E_2) \tag{5.4}$$
$$P(E_1\ or\ E_2) = P(E_1) + P(E_2) - P(E_1) \cdot P(E_2) \tag{5.5}$$

Hinweis: Wenn du die Verwendung von $+$ und $\cdot$ an die Symbolik der Booleschen Algebra aus MAT101 erinnert, ist das kein Zufall. So wie dort verwenden wir auch hier das Pluszeichen für *oder* (`OR`) und das Produkt für *und* (`AND`) und haben sogar ein *Komplement*: die Wahrscheinlichkeit, dass $E$ *nicht* eintrifft. Wir bezeichnen sie mit $P(not\ E)$.

Konkret bedeuten die Formeln (5.2) bis (5.5) somit:

1. Gehen wir davon aus, dass $E$ entweder eintreffen kann oder nicht (es aber keinen dritten Fall geben kann[7], dann ist die Summe aus $P(E) + P(not\ E)$ gleich 1.

2. Kennen wir die Wahrscheinlichkeit für das Eintreffen des Ereignisses $E$ und sie beträgt $P(E)$, so können wir aus Formel (5.2) unmittelbar Formel (5.3) ableiten und damit die Wahrscheinlichkeit angeben, dass $E$ *nicht* eintrifft: Wir rechnen uns einfach die Ergänzung auf 1 aus. Wir nennen das auch die **Gegenwahrscheinlichkeit**.

[6] Das sind wirklich nur die einfachsten Regeln. Es gibt dann zum Beispiel noch welche für eine *bedingte Wahrscheinlichkeit*, für die *totale Wahrscheinlichkeit*, oder für den so genannten *Satz von Bayes*, und noch vieles mehr. Für uns sind die obigen Regeln aber zunächst ausreichend.

[7] Ein «dritter Fall» wäre zum Beispiel *E trifft gleichzeitig ein und nicht*.

3. Kennen wir die Wahrscheinlichkeiten für das Eintreffen der Ereignisse $E_1$ und $E_2$ und betragen sie jeweils $P(E_1)$ und $P(E_2)$, so können wir mit Formel (5.4) die Wahrscheinlichkeit angeben, dass *sowohl* $E_1$ *als auch* $E_2$ eintreffen: Dazu multiplizieren wir die Einzelwahrscheinlichkeiten.

4. Kennen wir die Wahrscheinlichkeiten für das Eintreffen der Ereignisse $E_1$ und $E_2$ und betragen sie jeweils $P(E_1)$ und $P(E_2)$, so können wir mit Formel (5.5) die Wahrscheinlichkeit angeben, dass *entweder* $E_1$ *oder* $E_2$ eintreffen: Wir addieren dazu einfach die Einzelwahrscheinlichkeiten und ziehen davon die Wahrscheinlichkeit, dass $E_1$ *und* $E_2$ eintreffen wieder ab.

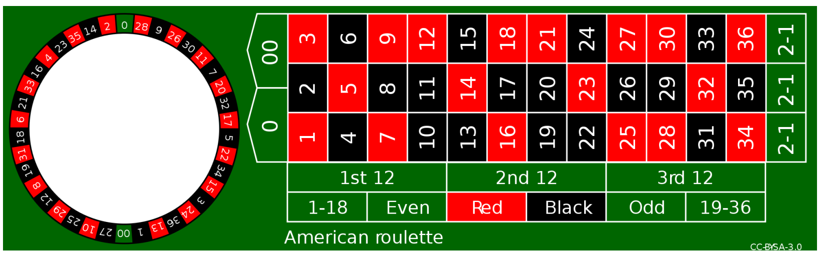
Alle genannten Regeln gelten für *unabhängige* Ereignisse und von solchen unabhängigen Ereignissen gehen wir hier aus. Das bedeutet, dass die Wahrscheinlichkeit für das Eintreffen von $E_1$ unabhängig davon ist, ob $E_2$ eingetroffen ist oder nicht und umgekehrt. Das ist zugegebenermaßen manchmal auf den ersten Blick nicht ganz offensichtlich. Würden heute die exakt selben sechs Lottozahlen gezogen werden wie bei der letzten Ziehung vorige Woche, würde das vermutlich viel Erstaunen auslösen (siehe S.102). Tatsächlich ist es aber völlig egal, welche Zahlen vorige Woche gezogen wurden. Die Lottomaschine hat kein «Gedächtnis», sondern geht nächste Woche wieder völlig unbedarft an ihre Arbeit.

**Beispiel 23** *Wir wissen: Beim Würfeln mit einem Würfel beträgt die Wahrscheinlichkeit, einen 6er zu würfeln,* $1/6$. *Wie groß ist die Wahrscheinlichkeit, keinen 6er zu würfeln?*

*Wenn wir schon wissen, dass die Wahrscheinlichkeit für einen 6er gleich* $1/6$ *ist, dann ist die Wahrscheinlichkeit, keinen 6er zu würfeln nach Formel 5.3 gleich*

$$P(not\ 6er) = 1 - P(6er) = 1 - 1/6 = 5/6 = 83.3\%$$

**Aufgabe 18** Beim American Roulette gibt es je 18 rote und schwarze Nummernfelder sowie zwei grüne Felder mit einer Null (siehe nachfolgende Abbildung[8]).




[8]Bildquelle: American roulette table and wheel layout. First published by Betzaar.com under the CreativeCommons Attribution+ShareAlike licence

Gib an:

1. Wie groß ist die Wahrscheinlichkeit auf eine «rote» Zahl?
2. Wie groß ist die Wahrscheinlichkeit auf eine Primzahl?
3. Wie groß ist die Wahrscheinlichkeit auf eine rote Primzahl?

**Beispiel 24** *Im einleitenden Beispiel auf Seite 97 haben wir uns gefragt, wie groß die Wahrscheinlichkeit ist, dass zwei Personen (aus einer Gruppe von 27) am selben Tag Geburtstag haben.*

*Mit der Definition (5.1) und den Formeln (5.3) und (5.4) können wir das nun angeben (wobei wir der Einfachheit halber Schaltjahre nicht berücksichtigen):*

*Bei zwei Personen beträgt die Wahrscheinlichkeit, dass sie* nicht *am selben Tag Geburtstag haben:* $\frac{364}{365} = 0.997$*. (Die zweite Person hat unter allen 365 Tagen des Jahres noch 364 Möglichkeiten, die nicht mit dem Tag der ersten Person übereinstimmen). Also ist die Wahrscheinlichkeit, dass sie* am selben Tag *Geburtstag haben die Gegenwahrscheinlichkeit* $(1 - 0.997) = 0.003$ *oder* $0.3\%$.

*Bei drei Personen darf die dritte weder mit der ersten noch mir der zweiten Person am gleichen Tag Geburtstag feiern. Diese Wahrscheinlichkeit ist* $\frac{364}{365} \cdot \frac{363}{365} = 0.992$*; die gesuchte Gegenwahrscheinlichkeit ist dann* $0.8\%$.

*Für 27 Personen gilt:* $\frac{364}{365} \cdot \frac{363}{365} \cdot \quad \cdots \quad \cdot \frac{339}{365} = 37\%$*. Das ist die Wahrscheinlichkeit, dass es unter den 27 Personen kein «Geburtstagspaar» gibt, bzw. umgekehrt: Die Wahrscheinlichkeit, dass unter 27 Personen zwei am selben Tag Geburtstag haben, beträgt* $63\%$.

*Nimmt man noch alle Ersatzspieler, die beiden Trainer und die beiden Schiedsrichterassistenten dazu, kommt man auf 41 Personen, die beim Champions League Finale 2012 am Fußballfeld herumliefen, und da ist die Wahrscheinlichkeit, dass zwei am selben Tag Geburtstag haben, bereits über* $90\%$*, wie sich leicht ausrechnen lässt.*

Vermutlich bist du ein wenig überrascht, dass bei 41 zufällig anwesenden Personen die Wahrscheinlichkeit, dass zwei an einem beliebigen aber selben Tag Geburtstag haben, so hoch ist – schließlich hat das Jahr ja 365 Tage. Aber du bist mit deiner Überraschung nicht alleine: Das Phänomen trägt auch den Namen «Geburtstagsparadoxon».

Beispiel 24 zeigt uns aber noch ein weiteres Dilemma: Während Formel 5.1 an sich ja ziemlich einfach ist, ist es manchmal schwierig bis unmöglich, die exakte Anzahl der Eintrittsfälle und Ausgangsmöglichkeiten herauszufinden. Im Geburtstagsbeispiel war es mit viel Nachdenken gerade noch möglich, aber Vieles, worüber wir Wahrscheinlichkeiten angeben wollen, ist überhaupt nicht zählbar. Wenn man sich zum Flughafen aufmacht, um eine Urlaubsreise zu beginnen, will man auf keinen Fall zu spät kommen und schätzt daher die Wahrscheinlichkeit, mit dem Auto in einen Stau zu kommen. Aber was sind da die Eintrittsfälle und die Ausgangsmöglichkeiten, die man ins Verhältnis zueinander setzen kann? Oder wenn für einen bestimmten Ort eine Regenwahrscheinlichkeit berechnet wird (vgl. Abb.5.1) – was wurde da abgezählt?

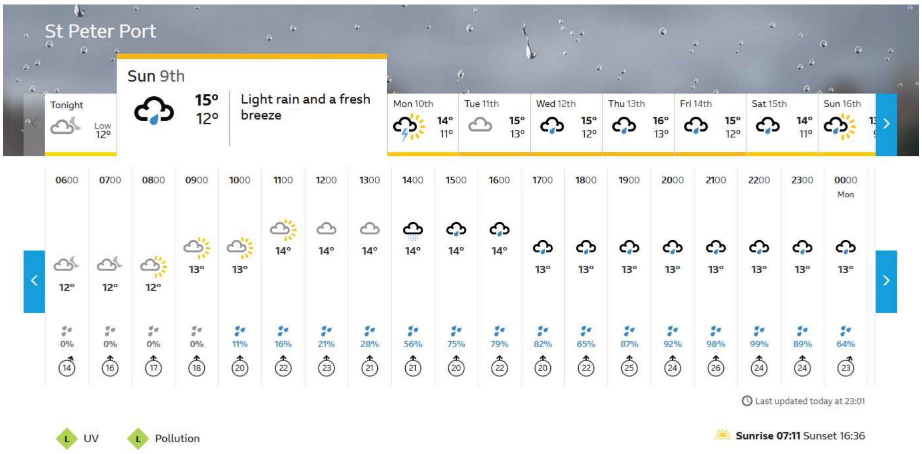



**Abb. 5.1:** Offenbar kann für jede Stunde eines bestimmten Tages eine Regenwahrscheinlichkeit angegeben werden. Aber wie kann man sie berechnen? (Quelle: www.bbc.com/weather)

Wir benötigen offenbar noch eine weitere Definition für die Wahrscheinlichkeit. Dazu kannst an dieser Stelle selbst ein kleines Experiment starten und eine Münze werfen und notieren, wie oft sie auf «Kopf» und wie oft auf «Zahl» fällt. Nach der klassischen Definition der Formel 5.1 beträgt die (theoretische) Wahrscheinlichkeit für «Kopf» $\frac{1}{2} = 0.5 = 50\%$. Das bedeutet aber nicht, dass wir im praktischen Experiment in genau 50% der Fälle Kopf erhalten, und schon gar nicht, dass wir dies genau bei jedem zweiten Wurf tun. Die Abfolge könnte zum Beispiel so aussehen (25 Würfe):

$$\{Z, Z, K, K, Z, Z, Z, Z, Z, Z, K, K, Z, K, Z, K, K, K, K, K, Z, Z, Z, Z, K\}$$

Im nächsten Schritt erinnern wir uns an die relative Häufigkeit $h$ aus Formel 2.5 (Seite 32). Wir erhalten sie, wenn wir die absolute Häufigkeit durch die Gesamtzahl der Elemente in der Stichprobe dividieren. Wir geben für unser Experiment die relative Häufigkeit für «Kopf» an:

Nach dem ersten Wurf $(Z)$ beträgt sie 0, ebenso nach den ersten beiden $(Z, Z)$. Nach dem dritten Wurf $(Z, Z, K)$ 0.33, nach dem vierten $(Z, Z, K, K)$ 0.5 etc. Nach dem 10. Wurf beträgt die relative Häufigkeit für «Kopf» 0.2, nach allen 25 Würfen 0.44. In Abb. 5.2 sind die relativen Häufigkeiten eingezeichnet.

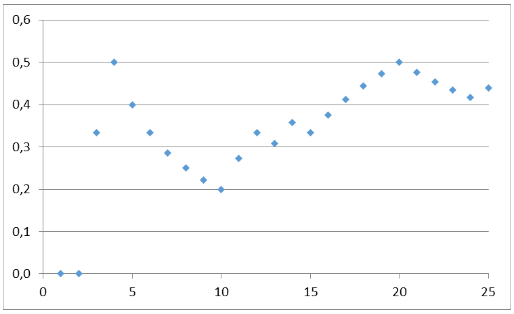



**Abb. 5.2:** Die relative Häufigkeit für «Kopf» bei 25maligem Münzwurf variiert in unserem Experiment anfänglich ziemlich stark.

Wir setzen das fort und werfen 50, 100, 500 und 1000 mal eine Münze und stellen die relativen Häufigkeiten wieder jeweils in einem Diagramm dar (Abb.5.3).

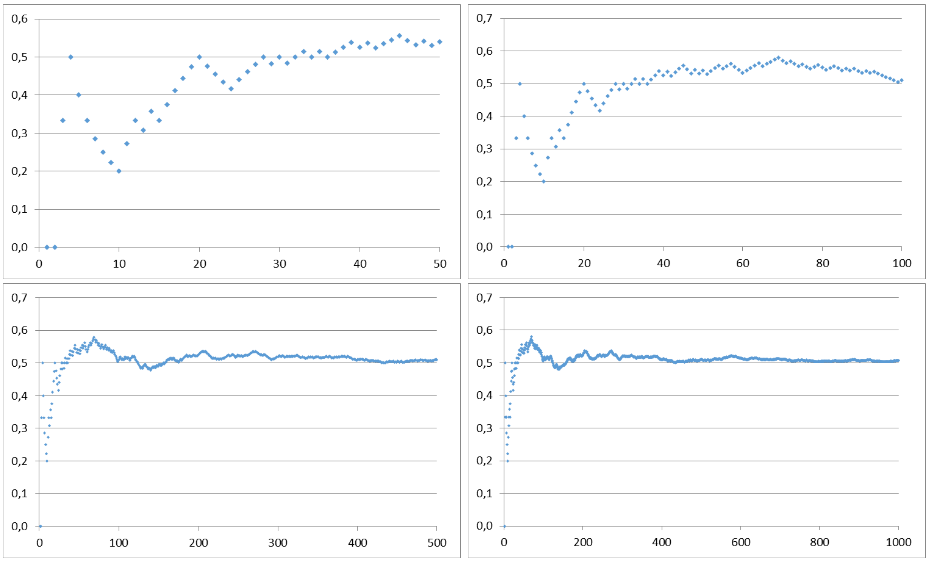



**Abb. 5.3:** Relative Häufigkeit für «Kopf» bei 50-/100-/500- und 1000maligem Münzwurf

Wir sehen: Je öfter wir die Münze werfen, desto eher nähert sich die relative Häufigkeit an die 50%-Linie an, also an die weiter oben (Beispiel 22) angegebene theoretische Wahrscheinlichkeit für das Auftreten von «Kopf».

Stellen wir uns jetzt vor, dass im Quotienten der Formel (2.5) das $n$ überhaupt unendlich groß wird, wir also unendlich viele Elemente in der Stichprobe (in unserem Beispiel: unendlich viele Münzwürfe) hätten. Mathematisch bedeutet dies, dass wir den *Grenzwert* der relativen Häufigkeit für «$n$ gegen $\infty$» angeben. Wir nennen diesen Grenzwert $P$, genauer $P(E)$:

$$\lim_{n \to \infty} \frac{k}{n} = P(E) \tag{5.6}$$

$P(E)$ ist eine Maßzahl für die Charakterisierung der Häufigkeit des Auftretens des Zufallsereignisses $E$ – die **Wahrscheinlichkeit** für $E$. In Worten ausgedrückt lautet die Definition 5.6:

*Mit wachsender Größe der Stichprobe konvergiert die relative Häufigkeit gegen die Wahrscheinlichkeit.*

Diese Definition der Wahrscheinlichkeit stammt von *Richard von Mises*[9].

Er beruft sich dabei auf das **Gesetz der großen Zahlen**, das besagt, dass die theoretische Wahrscheinlichkeit $P(E)$ umso besser geschätzt werden kann, je mehr unabhängige Ausführungen des Zufallsexperimentes durchgeführt werden. Oder anders ausgedrückt: Wenn man weiß, dass zum Beispiel beim Werfen einer Münze die Wahrscheinlichkeit auf «Kopf» 50% beträgt, heißt das bei 10 oder 20 Münzwürfen noch nicht viel. Aber umso öfter man die Münze wirft, desto eher wird sie tatsächlich in der Hälfte der Fälle auf «Kopf» fallen.

Für die absolute Häufigkeit $k$ gilt:

$$0 \leq k \leq n \tag{5.7}$$

und somit – wenn ich alles durch $n$ dividiere – für die relative Häufigkeit:

$$0 \leq \frac{k}{n} \leq 1 \tag{5.8}$$

Das gilt in jedem Fall, egal wie groß $n$ ist. Auch für $n \to \infty$ ändern sich diese Grenzen nicht, und daher gilt auch für die Wahrscheinlichkeit:

$$0 \leq P(E) \leq 1 \tag{5.9}$$

Die Wahrscheinlichkeit ist also eine reelle Zahl größer gleich Null und kleiner gleich Eins. Mit $P(E)$ können wir dann auf einer Skala von 0 bis 1 (bzw. 0% bis 100%) angeben, wie wahrscheinlich ein bestimmtes Ergebnis für zufällige Ereignisse und Experimente ist, bei denen es mehrere mögliche Ausgänge gibt. Dabei ist ein Ereignis, dem die Wahrscheinlichkeit 1 (bzw. 100%) zugeordnet ist, ein **sicheres Ereignis** ist, jenes mit der Wahrscheinlichkeit 0 (0%) ein **unmögliches Ereignis**. Umso näher $P(E)$ bei 1 liegt, desto wahrscheinlicher ist es, dass das Ereignis $E$ stattfindet; je näher es bei 0 liegt, desto größer ist die Wahrscheinlichkeit, dass es gar nicht eintrifft. Was aber nicht bedeutet, dass es unmöglich ist – schließlich gewinnt auch fast jede Woche jemand beim Lotto, obwohl die Wahrscheinlichkeit dafür nicht sehr hoch ist…

[9] österr.-amerikan. Mathematiker und Philosoph, 1883-1953

## 5.3 Die Wahrscheinlichkeitsverteilung von Zufallsgrößen

Sehen wir uns zunächst wieder ein Beispiel an: 30 Personen haben jeweils 100mal eine Münze geworfen und notiert, wie oft sie auf «Kopf» oder «Zahl» gefallen ist. Die Zufallsgröße, die wir in Folge näher betrachten wollen, ist die pro Person jeweils erzielte Anzahl des Ereignisses «Kopf». Gefühlsmäßig würden wir wohl erwarten, dass das bei den meisten Personen 50 Mal eintritt[10].

Das tatsächliche Ergebnis des Experiments ist in der Häufigkeitstabelle 5.1 gegeben.

| Anz. Kopf | k | h | Anz. Kopf | k | h | Anz. Kopf | k | h |
|---|---|---|---|---|---|---|---|---|
| 30 | 0 | 0 | 43 | 1 | 0.033 | 56 | 1 | 0.033 |
| 31 | 0 | 0 | 44 | 0 | 0 | 57 | 0 | 0 |
| 32 | 0 | 0 | 45 | 2 | 0.067 | 58 | 2 | 0.067 |
| 33 | 0 | 0 | 46 | 3 | 0.100 | 59 | 0 | 0 |
| 34 | 0 | 0 | 47 | 1 | 0.033 | 60 | 1 | 0.033 |
| 35 | 0 | 0 | 48 | 2 | 0.067 | 61 | 0 | 0 |
| 36 | 0 | 0 | 49 | 1 | 0.033 | 62 | 1 | 0.033 |
| 37 | 0 | 0 | 50 | 4 | 0.133 | 63 | 0 | 0 |
| 38 | 0 | 0 | 51 | 3 | 0.100 | 64 | 0 | 0 |
| 39 | 0 | 0 | 52 | 3 | 0.100 | 65 | 0 | 0 |
| 40 | 1 | 0.033 | 53 | 2 | 0.067 | 66 | 0 | 0 |
| 41 | 0 | 0 | 54 | 1 | 0.033 | 67 | 0 | 0 |
| 42 | 0 | 0 | 55 | 1 | 0.033 | 68 | 0 | 0 |

**Tabelle 5.1:** Absolute und relative Häufigkeiten eines Zufallsexperiments, bei dem 30 Spieler je 100 Mal eine Münze warfen. Die beobachtete Zufallsgröße ist dabei die Anzahl der «Kopf-Würfe». Kein Spieler hat in diesem Experiment weniger als 40 mal Kopf geworfen und keiner öfter als 62 mal.

In einem nächsten Schritt lassen wir das Experiment von 500 Personen durchführen, die wieder je 100mal eine Münze werfen. Das Ergebnis ist die Häufigkeitstabelle 5.2:

[10]Wobei dieses Gefühl auch trügerisch sein kann. *Bartovs et al.* haben 2023 in 350.757 Experimenten gezeigt, dass eine geworfene Münze eher dazu neigt, auf der gleichen Seite zu landen, auf der sie am Beginn des Wurfes gelegen ist. Konkret war das 178.079 mal der Fall, also in 51% der Würfe. Siehe: browse.arxiv.org/pdf/2310.04153v3.pdf

| Anz. Kopf | k | h | Anz. Kopf | k | h | Anz. Kopf | k | h |
|---|---|---|---|---|---|---|---|---|
| 30 | 0 | 0 | 43 | 13 | 0.026 | 56 | 17 | 0.034 |
| 31 | 1 | 0.002 | 44 | 20 | 0.04 | 57 | 17 | 0.034 |
| 32 | 0 | 0 | 45 | 27 | 0.054 | 58 | 12 | 0.024 |
| 33 | 2 | 0.004 | 46 | 28 | 0.056 | 59 | 3 | 0.006 |
| 34 | 0 | 0 | 47 | 37 | 0.074 | 60 | 3 | 0.006 |
| 35 | 2 | 0.004 | 48 | 27 | 0.054 | 61 | 1 | 0.002 |
| 36 | 0 | 0 | 49 | 36 | 0.072 | 62 | 1 | 0.002 |
| 37 | 0 | 0 | 50 | 41 | 0.082 | 63 | 1 | 0.002 |
| 38 | 5 | 0.01 | 51 | 40 | 0.08 | 64 | 2 | 0.004 |
| 39 | 3 | 0.006 | 52 | 42 | 0.084 | 65 | 1 | 0.002 |
| 40 | 4 | 0.008 | 53 | 43 | 0.086 | 66 | 0 | 0 |
| 41 | 6 | 0.012 | 54 | 29 | 0.058 | 67 | 0 | 0 |
| 42 | 13 | 0.026 | 55 | 23 | 0.046 | 68 | 0 | 0 |

**Tabelle 5.2:** Absolute und relative Häufigkeiten eines Zufallsexperiments, bei dem 500 Spieler je 100 Mal eine Münze warfen. Keine Person hat weniger als 31 oder öfter als 65 mal Kopf geworfen.

Schließlich lassen wir 10.000 Testpersonen Münzen werfen. Das Ergebnis sehen wir in Tabelle 5.3.

| Anz. Kopf | k | h | Anz. Kopf | k | h | Anz. Kopf | k | h |
|---|---|---|---|---|---|---|---|---|
| 30 | 0 | 0 | 43 | 301 | 0.0301 | 56 | 398 | 0.0398 |
| 31 | 2 | 0.0002 | 44 | 385 | 0.0385 | 57 | 321 | 0.0321 |
| 32 | 2 | 0.0002 | 45 | 471 | 0.0471 | 58 | 245 | 0.0245 |
| 33 | 5 | 0.0005 | 46 | 575 | 0.0575 | 59 | 168 | 0.0168 |
| 34 | 5 | 0.0005 | 47 | 632 | 0.0632 | 60 | 105 | 0.0105 |
| 35 | 12 | 0.0012 | 48 | 751 | 0.0751 | 61 | 67 | 0.0067 |
| 36 | 14 | 0.0014 | 49 | 818 | 0.0818 | 62 | 43 | 0.0043 |
| 37 | 32 | 0.0032 | 50 | 817 | 0.0817 | 63 | 32 | 0.0032 |
| 38 | 36 | 0.0036 | 51 | 776 | 0.0776 | 64 | 15 | 0.0015 |
| 39 | 77 | 0.0077 | 52 | 750 | 0.075 | 65 | 8 | 0.0008 |
| 40 | 110 | 0.011 | 53 | 647 | 0.0647 | 66 | 6 | 0.0006 |
| 41 | 158 | 0.0158 | 54 | 531 | 0.0531 | 67 | 5 | 0.0005 |
| 42 | 222 | 0.0222 | 55 | 458 | 0.0458 | 68 | 0 | 0 |

**Tabelle 5.3:** Absolute und relative Häufigkeiten eines Zufallsexperiments, bei dem 10 000 Spieler je 100 Mal eine Münze warfen. Minimum an «Kopf-Würfen»: 31, Maximum: 67

Die relativen Häufigkeiten können wir auch in einem Diagramm darstellen (Abb.5.4):

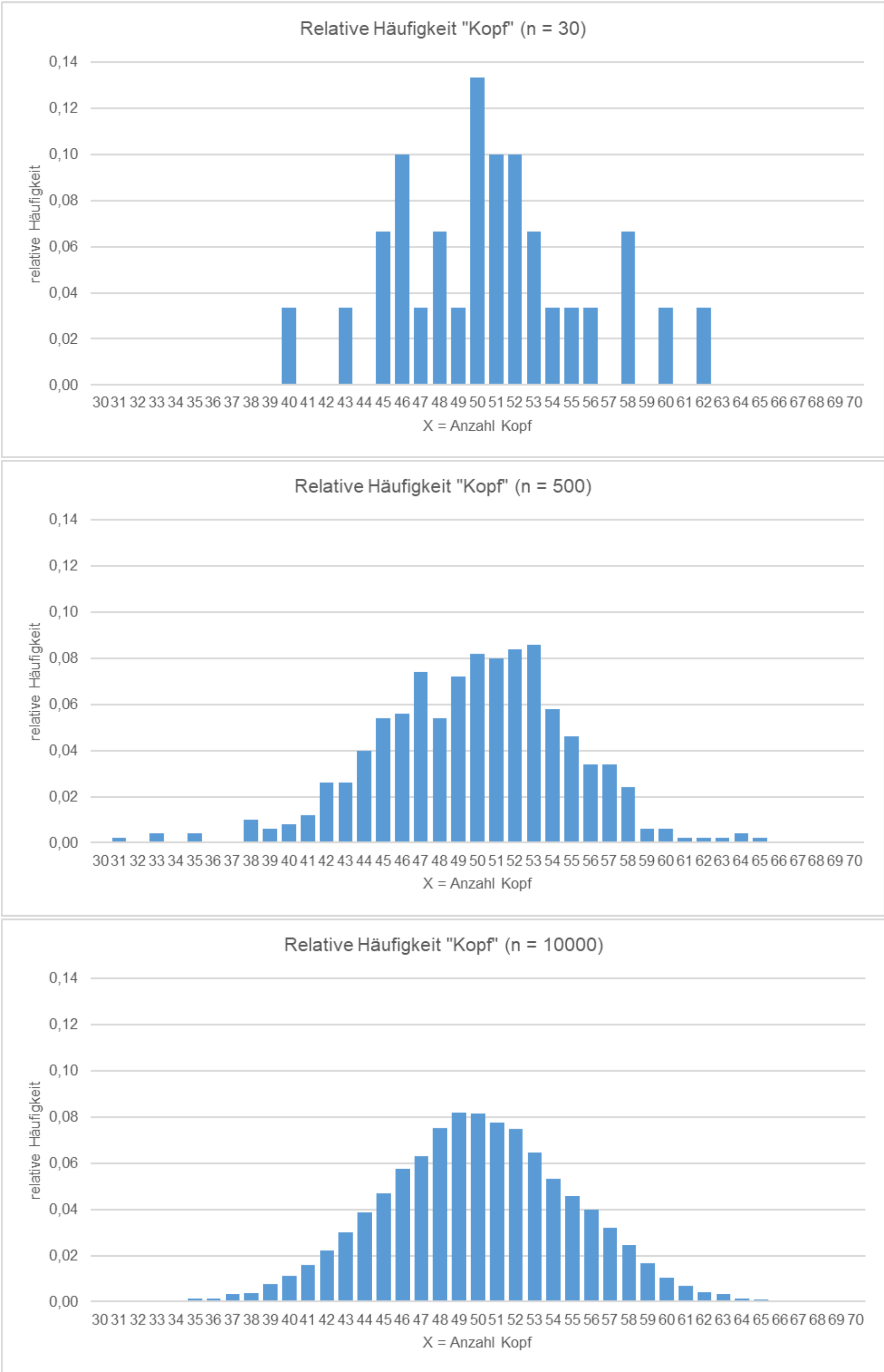



**Abb. 5.4:** Relative Häufigkeiten zu den Daten aus den Tabellen 5.1 (oben), 5.2 (Mitte) und 5.3 (unten)

Fassen wir noch einmal zusammen: Wir haben 30 Personen 100mal eine Münze werfen lassen; die Zufallsgröße, die wir dabei beobachten, ist die Anzahl, wie oft bei diesen 100 Würfen die Münze auf «Kopf» fällt. Wir erhalten 30 solche Zahlen (von jedem Mitspieler eine Anzahl, wie oft «Kopf» gefallen ist); der Umfang der Stichprobe ist also $n = 30$. Wir können die 30 Stichprobenelemente in einer Häufigkeitstabelle eintragen; eine Klasseneinteilung verwenden wir dabei nicht. Uns interessiert vor allem die relative Häufigkeit, außerdem stellen wir die Häufigkeitsverteilung in einem Säulendiagramm dar. Anschließend haben wir dasselbe Experiment von 500 Personen durchführen lassen und daraus eine neue Stichprobe (mit $n = 500$) erhalten und ausgewertet, und schließlich noch von 10 000 Personen eine Stichprobe mit $n = 10\,000$.

Beim Betrachten der drei Säulendiagramme in Abb.5.4 könnte Folgendes auffallen: Während bei einer Stichprobe aus $n = 30$ Elementen die Verteilung der Anzahl der Fälle, in denen eine 100mal zufällig geworfene Münze auf «Kopf» fiel, relativ willkürlich (und *zufällig*) aussieht, könnte man das Gefühl haben, dass die Diagramme von oben nach unten immer «kompakter» werden und im unteren Fall ($n = 10\,000$) die Verteilung eigentlich schon gar nicht mehr zufällig ausschaut, sondern so, als ob da irgendein Muster dahinter stecken würde. Mathematiker:innen[11] haben versucht, dieses Muster bzw. die dahinterliegende Mathematik zu finden, und zwar für den Fall, der die Häufigkeitsverteilung für den Fall $n = \infty$ beschreibt. Das Ergebnis sieht so aus (Abb. 5.5):

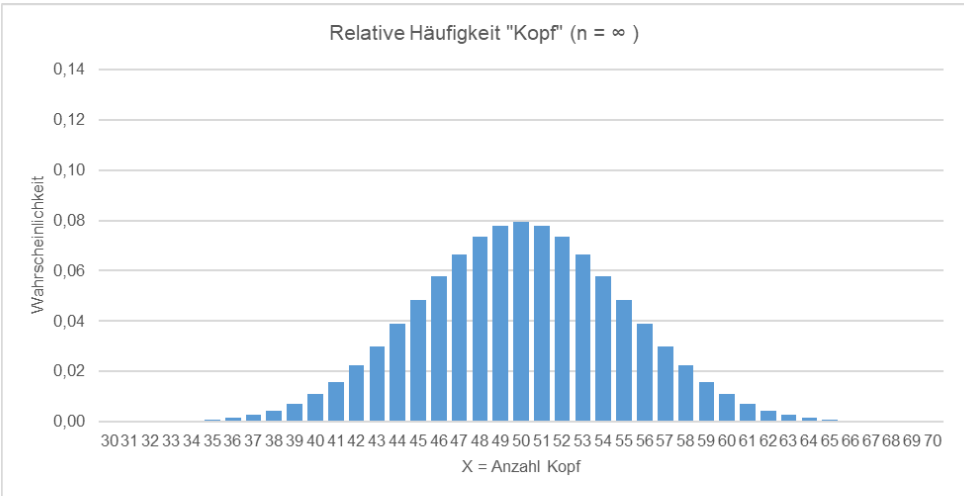



**Abb. 5.5:** Relative Häufigkeit der Zufallsgröße «Anzahl Kopf» bei je 100 Münzwürfen unendlich vieler Personen

[11] Zum Beispiel der Schweizer Mathematiker und Physiker *Jakob Bernoulli* (1655 - 1705). In seinem Buch *Ars conjectandi* (lat. für «*Die Kunst der Mutmaßung*»), posthum erschienen 1713, hat er den Grundstein für die oben beschriebene Wahrscheinlichkeitsfunktion gelegt.

In Tabelle 5.4 sind einige Zahlenwerte des in Abb.5.5 visualisierten Modells tabellarisch angegeben.

| $x$ | $P(X=x)$ | $x$ | $P(X=x)$ | $x$ | $P(X=x)$ | $x$ | $P(X=x)$ |
|---|---|---|---|---|---|---|---|
| 30 | 0.0000 | 40 | 0.0108 | 50 | 0.0796 | 60 | 0.0108 |
| 31 | 0.0001 | 41 | 0.0159 | 51 | 0.0780 | 61 | 0.0071 |
| 32 | 0.0001 | 42 | 0.0223 | 52 | 0.0735 | 62 | 0.0045 |
| 33 | 0.0002 | 43 | 0.0301 | 53 | 0.0666 | 63 | 0.0027 |
| 34 | 0.0005 | 44 | 0.0390 | 54 | 0.0580 | 64 | 0.0016 |
| 35 | 0.0009 | 45 | 0.0485 | 55 | 0.0485 | 65 | 0.0009 |
| 36 | 0.0016 | 46 | 0.0580 | 56 | 0.0390 | 66 | 0.0005 |
| 37 | 0.0027 | 47 | 0.0666 | 57 | 0.0301 | 67 | 0.0002 |
| 38 | 0.0045 | 48 | 0.0735 | 58 | 0.0223 | 68 | 0.0001 |
| 39 | 0.0071 | 49 | 0.0780 | 59 | 0.0159 | 69 | 0.0001 |

**Tabelle 5.4:** Wahrscheinlichkeitsverteilung der Zufallsgröße *«Anzahl Kopf» bei 100 Münzwürfen. Es handelt sich dabei um die auf* Jacob Bernoulli *(siehe Fußnote 11, S.111) zurückgehende* Binomialverteilung, *auf die wir auf Seite 113 noch zurückkommen werden.*

An dieser Stelle noch ein Geständnis: Wir haben keine 10 000 Menschen gefunden, die für uns für die Daten der Tabelle 5.3 $10\,000 \times 100 = 1\,000\,000$-mal gewürfelt haben – nicht einmal 30. Zum Glück können wir uns heute mit einer Computersimulation weiterhelfen. Das geht für 1 Million Realisierungen einer Zufallsvariable sogar noch in Excel.

Die oben verwendete Verteilung ist nur eines von mehreren möglichen Modellen, das wir verwenden können, um stochastische Phänomene in der Welt zu beschreiben. Schauen wir uns noch einige weitere an:

## 5.4 Die Modellierung der Verteilung diskreter Zufallsgrößen

Für *diskrete* Zufallsvariable können wir ein Modell angeben, das für jede einzelne mögliche Realisierung die Wahrscheinlichkeit für ihr Auftreten angibt. Dazu ordnen wir jeder Realisierung der Zufallsvariablen $x$ eine Wahrscheinlichkeit $P(x_i)$ zu. Die Funktion, die das leistet, ist die **Wahrscheinlichkeitsfunktion**:

$$f(x) = P(X = x) = \begin{cases} p_i & \text{für } x = x_i \\ 0 & \text{sonst} \end{cases} \tag{5.10}$$

Eine wichtige Frage ist manchmal auch die nach der Wahrscheinlichkeit, dass die Zufallsgröße $X$ kleiner oder gleich einer vorgegebenen Zahl $x$ ist oder zwischen zwei vorgegebenen Werten $a$ und $b$ liegt. Diese Fragen können mit Hilfe der *Verteilungsfunktion* beantwortet werden. Sie ist so definiert:

Der Funktionswert der **Verteilungsfunktion** $F(x)$ an der Stelle $x$ gibt die Wahrscheinlichkeit an, dass $X$ kleiner oder gleich $x$ ist. Im diskreten Fall entspricht das der Summe der Einzelwahrscheinlichkeiten:

$$F(x) = P(X \leq x) = \sum_{i=1}^{k} P(x = x_i) = \sum_{i=1}^{k} f(x_i) \text{ für } x_i \leq x \tag{5.11}$$

Die Verteilungsfunktion ist also das modellhafte Pendant zur kumulierten relativen Häufigkeit bei einer empirischen Häufigkeitsverteilung (vgl. S.33).

Formel 5.10 und 5.11 sind in dieser Form noch sehr abstrakte Definitionen – was konkret aber wird für die einzelnen $p_i$ eingesetzt? Dafür gibt es mehrere Möglichkeiten, von denen wir zwei näher anschauen: Die diskrete *Gleichverteilung* und die *Binomialverteilung*.

### Die Binomialverteilung

haben wir bereits in unserem «Münzwurf-Beispiel» (S.108ff.) kennen gelernt. Es handelt sich dabei um das mögliche Modell der Verteilung einer diskreten Zufallsvariable, und zwar für folgenden Fall eines Zufallsexperiments:

- ▷ Das Experiment besteht aus $N$ voneinander unabhängigen Versuchen.
- ▷ Bei jedem Versuch gibt es nur zwei mögliche Ausgänge. Einen bezeichnen wir – im Sinne des Experiments – als Erfolg, den anderen als Misserfolg
- ▷ Die Wahrscheinlichkeit für einen Erfolg beträgt $p$, für einen Misserfolg $(1 - p)$.
- ▷ Diese Wahrscheinlichkeiten $p$ und $(1 - p)$ bleiben bei jedem der $N$ Versuche gleich.

Für unser Beispiel war das: Es gab $N = 100$ Münzwürfe. Bei jedem Wurf gibt es zwei mögliche Ausgänge: Kopf oder Zahl. Die Wahrscheinlichkeit für Kopf beträgt $p = 0.5$ und für Zahl $(1 - p) = 0.5$. Das bleibt bei jedem der 100 Münzwürfe gleich. Die einzelnen Würfe sind auch unabhängig voneinander, d.h. die Münze hat kein «Gedächtnis», das z.B. nach 10 mal Kopf sagt: «Hey, jetzt kommt mal Zahl an die Reihe», sondern die Chancen sind auch nach 10 mal Kopf 50 : 50 für Kopf oder Zahl.

Die Binomialverteilung hat die *Wahrscheinlichkeitsfunktion*

$$f(x) = P(X = x) = \begin{cases} \binom{N}{x} p^x (1-p)^{N-x} & \text{für } x = 0, 1, \ldots, N \\ 0 & \text{sonst} \end{cases} \tag{5.12}$$

und die *Verteilungsfunktion*

$$F(x) = P(X \leq x) = \sum_{k=0}^{x} \binom{N}{k} p^k (1-p)^{N-k} \tag{5.13}$$

Wir brauchen diese Formeln nicht auswendig zu können – sie sind hier der Vollständigkeit halber angegeben (damit man z.B. die Werte der Tabelle 5.4 nachvollziehen kann). In unserem Beispiel ist $N = 100$ (weil jeder Teilnehmer 100 mal würfelt), $p = 0.5$ (weil die Wahrscheinlichkeit, dass «Kopf» kommt, 50% beträgt) und für $x$ werden der Reihe nach die $x$ aus der Tabelle 5.4 eingesetzt. Außerdem sehen wir an der Formel, warum die Verteilung «Binomialverteilung» heißt: Der dabei auftretende Koeffizient $\binom{N}{x}$ ist der so genannte *Binomialkoeffizient* (siehe Kap. 1.3 aus MAT101).

In *MS Excel* erhalten wir die Werte der Wahrscheinlichkeitsfunktion der Binomialverteilung mit `=BINOM.VERT(x;N;p;FALSCH)` und die Verteilungsfunktion mit `=BINOM.VERT(x;N;p;WAHR)`.
In *R* lautet der Befehl für die binomiale Wahrscheinlichkeitsfunktion `dbinom(x,N,p)` und `pbinom` für die Verteilungsfunktion.

Werte der Wahrscheinlichkeitsfunktion der Binomialverteilung in *Excel* und *R*

**Beispiel 25** *In einer Firma sind täglich zehn Server in Betrieb. Die Wahrscheinlichkeit, dass ein Server ausfällt, beträgt 1%. Durch redundante Serverkomponenten kann «normal» weitergearbeitet werden, solange nicht mehr als zwei Server gleichzeitig ausfallen. Wie hoch ist die Wahrscheinlichkeit für einen ausfallsicheren Betrieb?*

*Das Beispiel erfüllt alle Voraussetzungen einer Binomialverteilung: Die «Versuche» entsprechen den $N = 10$ Servern. Für jeden Server gibt es genau zwei Möglichkeiten: Er fällt aus (mit der Wahrscheinlichkeit $p = 0.01$), oder nicht (mit der Wahrscheinlichkeit $(1-p) = 0.99$). Und wir gehen davon aus, dass die Server unabhängig voneinander sind, d.h. auch evtl. Ausfälle unabhängig voneinander sind.*

*Damit ein ungestörtes Arbeiten möglich ist, dürfen nicht mehr als zwei Server ausfallen, d.h. es könnten $0, 1$ oder $2$ Server ausfallen.*

*Wir benötigen zunächst die Werte der Wahrscheinlichkeitsfunktion an den Stellen $0, 1$ und $2$, die wir z.B. aus Excel erhalten:* `BINOM.VERT(0;10;0,01;FALSCH)` *ergibt $90.44\%$, für $x = 1$ erhalten wir $9.14\%$ und für $x = 2$: $0.42\%$. In Summe sind das*

$$F(2) = P(X \leq 2) = 0.9044 + 0.0914 + 0.0042 = 0.9999$$

*(Wir hätten mit* `BINOM.VERT(2;10;0,01;WAHR)` *auch direkt den Wert der Verteilungsfunktion an der Stelle 2 rechnen können und wären auf denselben Wert gekommen).*

*D.h. mit einer Wahrscheinlichkeit von 99.99% fallen nicht mehr als 2 Server aus und es kann in der Firma ungestört gearbeitet werden – also zumindest was die Funktionalität der Server betrifft. . .*

**Aufgabe 19** Ein Statistiktest besteht aus 6 Single-Choice Fragen mit jeweils 4 Antwortmöglichkeiten. (Single-Choice = genau eine Antwort aus den 4 möglichen ist richtig). Für jemanden, der sich nicht auf den Test vorbereitet hat und nach Belieben zufällige Antworten ankreuzt, beträgt die Erfolgswahrscheinlichkeit pro Frage $p = 25\%$. Wie groß ist die Wahrscheinlichkeit, dass diese Person den Test positiv besteht, d.h. mindestens 3 Fragen richtig beantwortet?

### Die diskrete Gleichverteilung

Eine andere, einfache Verteilung, die eine Zufallsgröße haben kann, ist die diskrete **Gleichverteilung**. Sie ordnet allen innerhalb des Intervalls $[a, b]$ liegenden Werten einer Zufallsgröße die gleiche Wahrscheinlichkeit zu.

Die *Wahrscheinlichkeitsfunktion* der diskreten Gleichverteilung lautet:

$$f(x) = \begin{cases} \frac{1}{k} & \text{für alle } x_i \text{ mit } i = 1, \ldots, k \\ 0 & \text{sonst} \end{cases} \tag{5.14}$$

**Beispiel 26** *Beim Würfeln mit einem «unverfälschten» Würfel gibt es sechs mögliche Ausgänge: Man kann $1, 2, 3, 4, 5$ oder $6$ würfeln. Im Sinne der Formel (5.14) können wir also schreiben: $k = 6$ und $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$, $x_5 = 5$, $x_6 = 6$.*

*Jetzt können wir den Wert der Wahrscheinlichkeitsfunktion zum Beispiel an der Stelle $x = 2$ ausrechnen:*

$$f(2) = \frac{1}{6} = 0.167$$

*oder an der Stelle 7:*

$$f(7) = 0$$

*Das stimmt auch mit unserer bisherigen Vorstellung überein: Die Wahrscheinlichkeit, einen 2er zu würfeln, beträgt 16.7%, die Wahrscheinlichkeit, einen 7er zu würfeln ist hingegen gleich Null. (Jetzt wissen wir auch, warum Mathematiker in den Formeln 5.10 und 5.14 die «sonst»-Zeile eingefügt haben).*

**Aufgabe 20** Auf Seite 52 haben wir den Begriff Modalwert kennengelernt und ihn u.a. auch als als *wahrscheinlichsten Wert* bezeichnet. Was ist der wahrscheinlichste Wert der in Abb.5.5 dargestellten Verteilung?

**Aufgabe 21** Hat eine Gleichverteilung auch einen Modalwert?

## 5.5 Die Modellierung der Verteilung stetiger Zufallsgrößen

Bei stetigen Zufallsgrößen müssen wir Folgendes beachten: Die Anzahl aller möglichen Realisierungen einer stetigen Zufallsvariable ist nicht abzählbar sondern unendlich groß – so wurden ja stetige Variable auf Seite 13 definiert. Wenn wir aber in Formel 5.1 im Nenner $\infty$ einsetzen, erhalten wir: $P(E) = 0$. Das bedeutet, dass wir im stetigen Fall einem bestimmten $x$ keine Wahrscheinlichkeit $P(X = x)$ zuordnen können! Es geht aber für ein Intervall $[a, b]$.

Beginnen wir mit der **Verteilungsfunktion** $F(x)$. Formal sieht sie so ähnlich aus wie 5.11, allerdings wird die Summe durch ein Integral ersetzt:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt \tag{5.15}$$

Die in Formel 5.15 auftretende Funktion $f(t)$ nennen wir **Dichtefunktion** der Verteilung (auch: *Wahrscheinlichkeitsdichte* bzw. nur *Dichte*, im Englischen: *probability density function*, abgekürzt PDF). Sie ist so was ähnliches wie die Wahrscheinlichkeitsfunktion im diskreten Fall. Mit dem Unterschied, dass wir – wie bereits oben begründet – nicht die Wahrscheinlichkeit für ein bestimmtes $x$ angeben können, sondern nur für ein Intervall. Abb. 5.6 zeigt ein Beispiel für eine Dichtefunktion und den Zusammenhang zur Verteilungsfunktion $F(x)$. Wir sehen dabei unter anderem: Die Wahrscheinlichkeit $P(a < X \leq b)$ entspricht der

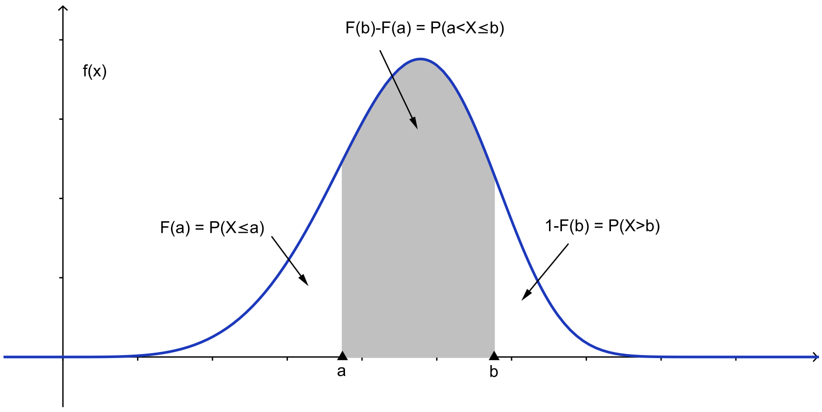



**Abb. 5.6:** Eine mögliche Dichtefunktion $f(x)$ einer stetigen Zufallsgröße und der Zusammenhang zur Verteilungsfunktion $F(x)$

Fläche zwischen der $x$-Achse und der Dichtefunktion $f(x)$ von $a$ bis einschließlich $b$.

Wie aus Abb.5.6 ersichtlich, gelten für die Verteilungsfunktion folgende wichtige Zusammenhänge:

$$P(X \leq a) = F(a) = \int_{-\infty}^{a} f(x)dx \tag{5.16}$$

$$P(X > b) = 1 - F(b) = \int_{b}^{+\infty} f(x)dx \tag{5.17}$$

$$P(a < X \leq b) = F(b) - F(a) = \int_{a}^{b} f(x)dx \tag{5.18}$$

Es gibt eine große Anzahl von Modellen für die Verteilung stetiger Zufallsgrößen. Abb.5.7 zeigt ein paar Beispiele für häufig vorkommende Dichtefunktionen. Konkret werden wir uns zwei näher ansehen: Die stetige Gleichverteilung und die Normalverteilung.

### Die stetige Gleichverteilung

ist ähnlich definiert wie die diskrete Gleichverteilung: Sie ordnet allen gleichgroßen Intervallen aus dem Bereich von $a$ bis $b$ die gleiche Wahrscheinlichkeit $\neq 0$ zu, außerhalb dieses Bereichs ist die Wahrscheinlichkeit gleich Null. Der

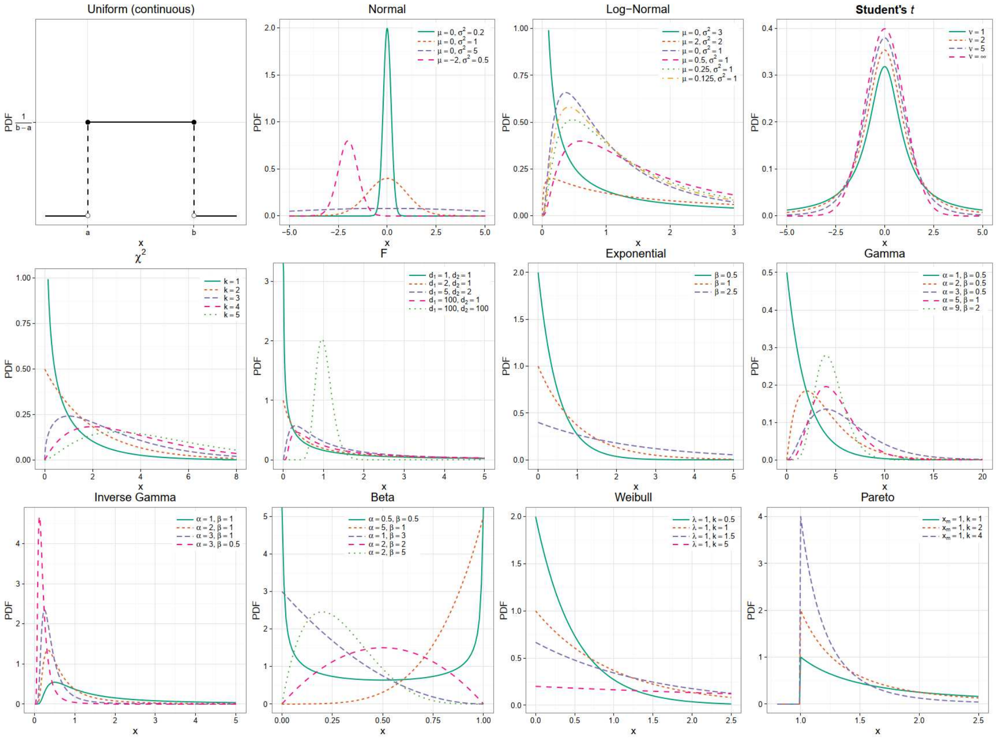



**Abb. 5.7:** Darstellung der Dichtefunktionen einiger stetiger Wahrscheinlichkeitverteilungen. (Quelle: Vallentin, Matthias. 2017. The Probability and Statistics Cookbook. https://github.com/mavam/stat-cookbook. CC BY-NC-SA 4.0)

Unterschied zur diskreten Gleichverteilung ist der, dass im Bereich $[a, b]$ nicht eine abzählbare Anzahl von Werten liegt, sondern unendlich viele Intervalle mit unendlich vielen Zahlen aus $\mathbb{R}$.

Die stetige Gleichverteilung hat die *Dichtefunktion*

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq x \leq b \\ 0 & \text{sonst} \end{cases} \tag{5.19}$$

Die *Verteilungsfunktion* der stetigen Gleichverteilung im Bereich $a \leq x \leq b$ ist gegeben durch:

$$P(X \leq x) = F(x) = \frac{x-a}{b-a} \tag{5.20}$$

Der Graph der Dichtefunktion hat ein rechteckiges Aussehen. Dichte- und Verteilungsfunktion sind in Abb.5.8 dargestellt und wir sehen darin auch: Wenn $x < a$ oder $x > b$, dann ist $F(x) = 0$.

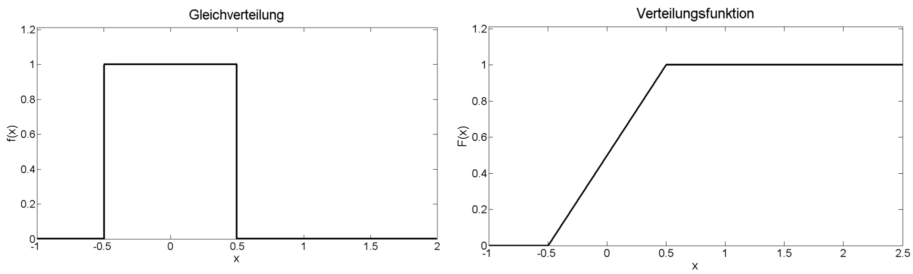



**Abb. 5.8:** Dichtefunktion und Verteilungsfunktion der stetigen Gleichverteilung in den Grenzen $-0.5 \leq x \leq 0.5$.

**Beispiel 27** *Wie groß ist die Wahrscheinlichkeit, dass eine gleichverteilte stetige Zufallsgröße zwischen den Werten $x_1$ und $x_2$ liegt?*

*Nach Formel 5.18 können wir angeben:*

$$P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1)$$

*und hier jetzt die konkreten Funktion aus Formel 5.20 einsetzen:*

$$P(x_1 \leq X \leq x_2) = \frac{x_2 - a}{b - a} - \frac{x_1 - a}{b - a} = \frac{x_2 - a - x_1 + a}{b - a} = \underline{\underline{\frac{x_2 - x_1}{b - a}}}$$

**Aufgabe 22** Am Bahnhof Meidling wird folgende Information durchgesagt: «Der Zug Richtung Wiener Neustadt hat zwischen 5 und 15 Minuten Verspätung». Angenommen es gibt keinen Grund zur Annahme, dass die Verspätung eher bei 5 oder 15 Minuten liegt, sondern der Zug tatsächlich «irgendwann» in diesem Intervall eintreffen wird: Wie groß ist die Wahrscheinlichkeit, dass die Verspätung maximal 7 Minuten ausmacht?

### Die Normalverteilung

Eine für uns sehr wichtige Verteilung von Daten ist die **Normalverteilung**, auch *Gaußsche Verteilung* genannt[12]. Viele Sachverhalte aus den Natur- und Sozialwissenschaften sind annähernd normalverteilt (oder zumindest modellieren wir sie so). In den modernen Wirtschaftswissenschaften folgen die Phänomene zwar manchmal anderen Mustern, es lässt sich aber zeigen, dass immer dann, wenn wir eine entsprechend große Anzahl zufälliger Ereignisse mit einer beliebigen Verteilung zu einer einzigen Zufallsvariable zusammenfassen können, diese (zumindest annähernd) normalverteilt ist[13].

[12] *Johann Friedrich Carl Gauß*, deutscher Mathematiker und Geodät, 1777-1855

[13] Das besagt der so genannte *Zentrale Grenzwertsatz*, auf den wir hier aber nicht näher eingehen.

Die mathematische Funktion der Normalverteilung ist ziemlich kompliziert und für uns nicht weiter wichtig. Wir werden einfach ein Computerprogramm verwenden, wenn wir tatsächlich einen Wert ausrechnen müssen. Wir merken uns nur, dass die Normalverteilung über zwei Parameter definiert ist: den theoretischen Mittelwert $\mu$ (der **Erwartungswert** genannt wird) und die theoretische Varianz $\sigma^2$. Wie wir zu diesen beiden Parametern kommen, werden wir uns im nächsten Kapitel genauer ansehen. Einstweilen setzen wir, wenn uns die theoretischen Modellparameter nicht bekannt sind, für den Erwartungswert das arithmetische Mittel ein und für die theoretische Varianz den Wert der empirischen Varianz.

Zwei Datensätze, die wir untersuchen, haben üblicherweise – auch wenn sie beide normalverteilt sind – unterschiedliche Erwartungswerte und unterschiedliche Varianzen, und somit eine andere Form der Normalverteilung. Letztlich gibt es unzählig viele Normalverteilungen; einige sind beispielhaft in Abb.5.9 dargestellt. Wenn wir unsere normalverteilten Daten wie auf Seite 74 angegeben *standardisieren* (zu *z-Werten* mit Mittelwert 0 und Standardabweichung 1), so sind diese **standardnormalverteilt**, d.h. verteilt nach einer Normalverteilung mit Erwartungswert $\mu = 0$ und $\sigma^2 = 1$.

Grafisch ähnelt die Dichtefunktion der Normalverteilung der Form einer «Glocke» und wird daher auch *Glockenkurve* (auch: *Gaußsche Glockenkurve*) genannt (Abb.5.9).

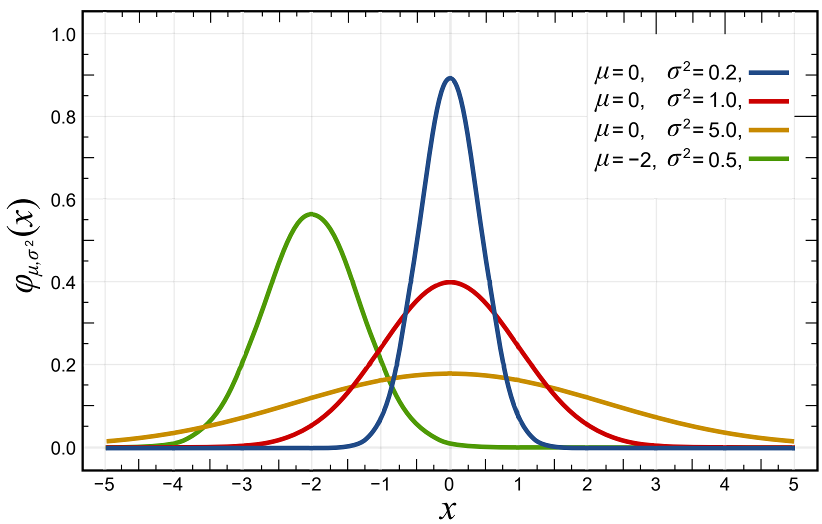



**Abb. 5.9:** Gaußsche Glockenkurven: Verschiedene Dichtefunktionen zur Normalverteilung mit unterschiedlichen Erwartungswerten und Varianzen. Der Scheitel der jeweiligen Kurven liegt bei $x = \mu$, ihre Wendepunkte im Abstand $\pm\sigma$ vom Scheitelwert. Die Kurve ist umso höher und steiler, je kleiner $\sigma$ ist. (Quelle: https://commons.wikimedia.org/w/index.php?curid=3817954)

Wie wir aus der Abbildung sehen können, hat die Normalverteilung folgende Eigenschaften:

Der Parameter $\mu$ bestimmt das Zentrum der Kurve, $\sigma$ ihre «Schlankheit» (Breite). Der Scheitel der Dichtefunktion – also das Maximum – liegt bei $x = \mu$. Die Wendepunkte (das ist der mathematische Ausdruck für Punkte, wo die Kurve am «steilsten» ist) liegen im Abstand $\pm\sigma$ von $\mu$. Je kleiner $\sigma$ ist, desto schmäler ist die Glockenkurve, je größer $\sigma$ ist, desto breiter und flacher wird sie.

Die beiden Enden der Kurve kommen zwar sehr nahe an die $x$-Achse heran, berühren sie aber nie. (Mathematisch heißt das: Sie nähern sich *asymptotisch* an die *Abszisse* an). Und obwohl das so ist, hat die Fläche unter der Kurve überraschenderweise einen endlichen Wert, nämlich genau 1. Das gilt für alle einzelnen unterschiedlichen Kurven der Abb.5.9: Die Gesamtfläche unter der orangen, grünen, blauen und magenta-farbenen Kurve ist jeweils genau 1 Einheit groß.

Weiters ist erkennbar, dass die Normalverteilung eine um $\mu$ symmetrische Verteilung ist, d.h. betragsmäßig gleich große positive oder negative Abweichungen vom Erwartungswert sind gleich wahrscheinlich. Daher ist der Erwartungswert auch gleichzeitig der Median der Verteilung. Die Wahrscheinlichkeit für eine Abweichung vom Erwartungswert ist umso geringer, je größer diese Abweichung ist. Große Abweichungen sind also weniger wahrscheinlich als kleine.

Und: Für eine normalverteilte Zufallsgröße gilt:

- ▷ ca. 68% aller Daten liegen im Intervall $\mu \pm 1 \cdot \sigma$, also innerhalb eines Bereichs im Abstand von maximal einer Standardabweichung um den Erwartungswert.
- ▷ ca. 95% aller Daten liegen im Intervall $\mu \pm 2 \cdot \sigma$, also innerhalb eines Bereichs im Abstand von maximal zwei Standardabweichungen um den Erwartungswert.
- ▷ ca. 99.7% aller Daten liegen im Intervall $\mu \pm 3 \cdot \sigma$, also innerhalb eines Bereichs im Abstand von maximal drei Standardabweichungen um den Erwartungswert.

**Beispiel 28** *Angenommen im Studiengang WIBA beträgt das (langjährig beobachtete) Durchschnittsalter der Studierenden $\bar{x} = 37$ Jahre mit einer Standardabweichung von $s = 5$ Jahren. Wie groß ist die Wahrscheinlichkeit, dass unter der Annahme einer Normalverteilung ein:e (beliebig ausgewählte:r) Studierende:r nicht älter als 40 Jahre alt ist?*

*Allgemein finden wir die Antwort in Formel 5.16. Konkret suchen wir den Wert der Verteilungsfunktion an der Stelle $P(X \leq 40) = F(40)$ für eine Normalverteilung mit*

In *MS Excel* können wir für eine normalverteilte Zufallsgröße sowohl den Wert der Verteilungs- als auch den der Dichtefunktion ausrechnen:
Den Wert der *Dichtefunktion* einer normalverteilten Zufallsgröße an der Stelle $x$ erhalten wir aus
`=NORM.VERT(x;Erwartungswert;Standabweichung;FALSCH)`
Den Wert der *Verteilungsfunktion* einer normalverteilten Zufallsgröße an der Stelle $x$ erhalten wir aus
`=NORM.VERT(x;Erwartungswert;Standabweichung;WAHR)`

In *R* erhalten wir den Wert der Verteilungsfunktion der Normalverteilung mit `pnorm`, die Dichtefunktion mit `dnorm`.

*dem Erwartungswert 37 und der Standardabweichung 5. Laut EXCEL beträgt dieser Wert 0.7257 (`=NORM.VERT(40;37;5;WAHR)`).*

*Die Wahrscheinlichkeit, dass ein Studierender nicht älter als 40 Jahre ist, beträgt demnach 72.6%*

**Aufgabe 23** Angenommen im Studiengang WIBA beträgt das (langjährig beobachtete) Durchschnittsalter der Studierenden $\bar{x} = 37$ Jahre mit einer Standardabweichung von $s = 5$ Jahren. Wie groß ist die Wahrscheinlichkeit, dass unter der Annahme einer Normalverteilung ein:e (beliebig ausgewählte:r) Studierende:r zwischen 32 und 42 Jahre alt ist?

In *R* können wir einen Vektor von Zufallszahlen erzeugen, die einer bestimmten Verteilung folgen. Zum Beispiel erhalten wir 100 *gleichverteilte* Zufallszahlen im Intervall $(a, b)$ mit der Funktion `runif(100,a,b)`. *Binomialverteilte* Zufallszahlen erhalten wir mit dem Befehl `rbinom`, normalverteilte mit `rnorm`. Die genaue Syntax dieser Befehle ist in der R-Hilfe beschrieben (aufrufbar mit dem Befehl `help()`, also z.B. `help(rnorm)`).

Um überhaupt Hilfe über alle in *R* implementierten Verteilungen zu erhalten, gibt man den Befehl `help(Distributions)` ein.

## Zusammenfassung

Der eine oder die andere mag sich vielleicht fragen, wozu wir in der Statistik die Wahrscheinlichkeitsrechnung und diese manchmal ziemlich kompliziert anmutenden theoretischen Verteilungsfunktionen brauchen[14]. Dazu erinnern wir uns an das auf Seite 6 Gesagte über das Ziel, das wir mit der Anwendung der Statistik erreichen wollen: Wir wollen ***Modelle*** *finden, mit denen wir besser verstehen können, wie oder warum bestimmte Phänomene in der realen Welt funktionieren*. Die Wahrscheinlichkeitsverteilungen dieses Kapitels liefern uns für diese Modellbildung eine gute Basis. Wir versuchen so gut wie möglich, auf den ersten Blick «regellos» verteilten empirisch gewonnenen Daten eine Struktur zu geben, am besten eine mathematisch formalisierte Struktur. Nur so können wir letztlich die dahinterliegenden Phänomene erklären und sogar einen Nutzen für zukünftige Ereignisse ziehen, indem wir zum Beispiel Computer dazu bringen, das Verhalten von Menschen, Wirtschaftssystemen, dem Klima, Verkehrsströmen, technischen oder Produktions-Abläufen, oder die Ausbreitung von Virus-Epidemien etc. zu simulieren.

Dafür stehen uns eine Menge von Modellen zur Verfügung. Abb.5.10 zeigt eine Übersicht dazu – oder zumindest das, was Statistiker:innen als «Übersicht» bezeichnen. Wenn dir das ziemlich komplex und/oder chaotisch vorkommt, keine Angst: Wir werden uns in dieser Einführungslehrveranstaltung nicht mit mehr als der Normalverteilung auseinandersetzen. Und mit dem nächsten Kapitel gleich die Tür dorthin öffnen und das Feld der *schließenden Statistik* betreten.

Wer sich zuvor noch ein paar anschauliche Beispiele ansehen möchte, in denen Wahrscheinlichkeiten visualiert werden, kann das unter seeing-theory.brown.edu tun.

[14] An dieser Stelle können wir nicht auf die eingehen, die sich schon seit dem ersten Kapitel fragen, wozu wir das Ganze brauchen. . .

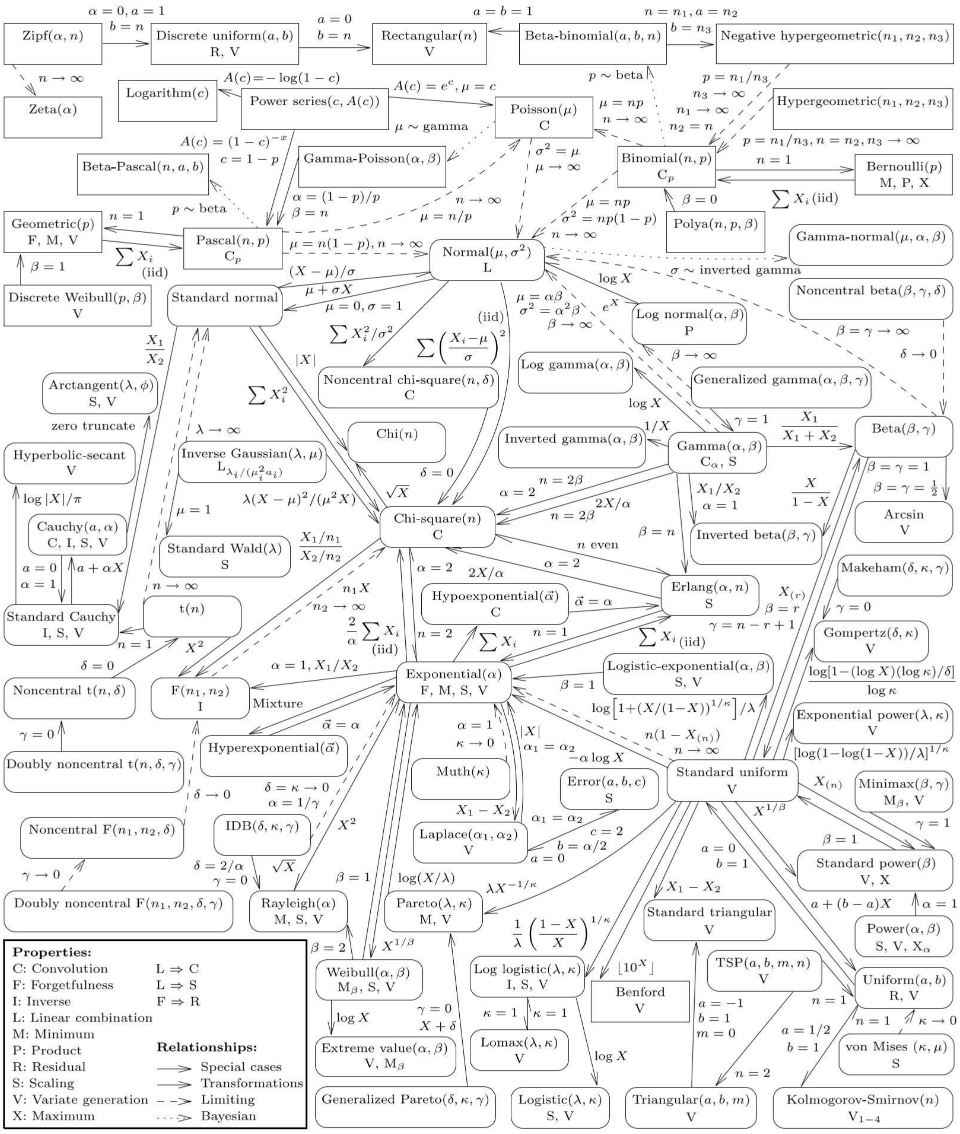



**Abb. 5.10:** Univariate Wahrscheinlichkeitsverteilungen und wie sie zusammenhängen. (Quelle: Leemis, Lawrence M. and McQueston, Jacquelyn T. 2008. “Univariate Distribution Relationships”. *The American Statistician* 62(1) p.47)